

UNITED STATES PATENT AND TRADEMARK OFFICE



UNITED STATES DEPARTMENT OF COMMERCE United States Patent and Trademark Office Address: COMMISSIONER FOR PATENTS P.O. Box 1450 Alexandria, Virginia 22313-1450 www.uspto.gov

APPLICATION N	О.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/680,517	09/680,517 10/06/2000		Hidehito Kubo	043034/0158	9120
22428	7590	12/01/2004		EXAMINER	
FOLEY SUITE 50	AND LA	RDNER	OPIE, GEORGE L		
	TREET N	W	ART UNIT	PAPER NUMBER	
WASHINGTON, DC 20007				2126	
				DATE MAILED: 12/01/2004	

Please find below and/or attached an Office communication concerning this application or proceeding.

	Application No.	Applicant(s)
Office Action Comment	09/680,517	Hidehito Kubo
Office Action Summary	Examiner	Art Unit
	George L. Opie	2126
The MAILING DATE of this communication a Period for Reply	ppears on the cover sheet with the	correspondence address
A SHORTENED STATUTORY PERIOD FOR REI		TH(S) FROM
 Extensions of time may be available under the provisions or after SIX (6) MONTHS from the mailing date of this comr If the period for reply specified above is less than thirty (30) be considered timely. If NO period for reply is specified above, the maximum state communication. Failure to reply within the set or extended period for reply w Status 	nunication. days, a reply within the statutory minimulutory period will apply and will expire SIX	m of thirty (30) days will (6) MONTHS from the mailing date of this
1) X Responsive to communication(s) filed on 3	0 August 2004	
	This action is non-final.	•
Since this application is in condition for allocation accordance with the practice und	- wance except for formal matters,	
Disposition of Claims		
4) X Claim(s) 1-49 is/are pending in the applica	tion.	
4a) Of the above claim(s) is/are withd	rawn from consideration.	
5) X Claim(s) 33-36, 40-43 and 46 is/are allowed	ed.	
6) X Claim(s) 1-8, 11-22, 25-32, 37-39, 44-45 ar	nd 47-49 is/are rejected.	
7) X Claim(s) <u>9-10, 23-24 and 31</u> is/are objected to	,	
8) Claim(s) are subject to restriction and	d/or election requirement.	
Application Papers		
9) The specification is objected to by the Exam	niner.	
10) The drawing(s) filed on is/are of	bjected to by the Examiner.	
11) The proposed drawing correction filed on	is: a) approved b)	_ disapproved.
12) The oath or declaration is objected to by the	e Examiner.	
Priority under 35 U.S.C. § 119		
13) X Acknowledgment is made of a claim for fore	eign priority under 35 U.S.C. § 11	9(a)-(d).
a) X All b) Some * c) None of the CEF	RTIFIED copies of the priority doc	uments have been:
1. X received.		
2 received in Application No. (Series C	ode / Serial Number)	
3 received in this National Stage applic	ation from the International Burea	au (PCT Rule 17.2(a)).
* See the attached detailed Office action for a l	ist of the certified copies not rece	ived.
14) Acknowledgement is made of a claim for	domestic priority under 35 U.S.C.	& 119(e).
Attachment(s)		
 14) X Notice of References Cited (PTO-892) 15) Notice of Draftsperson's Patent Drawing Review (PTO-948) 16) Information Disclosure Statement(s) (PTO-1449) Paper No 	18) Notice of Inform	mary (PTO-413) Paper No(s) nal Patent Application (PTO-152) cs for USP6,665,716 USP5,978,844

DETAILED ACTION

This Office Action is responsive to the Amendment filed 30 August 2004,in which claims 1-4, 7-18 and 21-46 were amended. New claims 47-49 were added. Hence, claims 1-49 are now pending.

1. Request for copy of Applicant's response on floppy disk:

Please help expedite the prosecution of this application by including, along with your amendment response in paper form, an electronic file copy in WordPerfect, Microsoft Word, or in ASCII text format on a 3½ inch IBM format floppy disk. Please include all pending claims along with your responsive remarks. Only the paper copy will be entered -- your floppy disk file will be considered a duplicate copy. Signatures are not required on the disk copy. The floppy disk copy is not mandatory; however, it will help expedite the processing of your application. Your cooperation is appreciated.

Allowable Subject Matter

- 2. Allowed Claims: 33-36, 40-43 and 46
- 3. Claims 9-10, 23-24 and 31 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and the intervening claims.
- 4. The U.S. Patents used in the art rejections below have been provided as text documents which correspond to the U.S. Patents. The relevant portions of the text documents are cited according to page and line numbers in the art rejections below. For the convenience of Applicant, the cited sections are highlighted in the *text documents*.
- 5. Claim Rejections 35 U.S.C. § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains.

Patentability shall not be negatived by the manner in which the invention was made.

6. Claims 1-8 11-22, 25-28, 37 and 47-49 are rejected under 35USC§103(a) as being unpatentable over the Admitted Prior Art (APA) in view of Sitbon et al. (U.S. Patent 5,951,634) and Hirata et al. (U.S. Patent 6,665,716).

As to claim 1, the APA teaches a load balancing method in a system ("Optimal Load Balancing in Distributed Computer Systems") comprising a plurality of computers (servers) for processing transaction processing requests ("distributing ... messages each requiring small-scale processing") originating from a plurality of terminals, comprising steps of:

- a) estimating load states of respective ones of the computers ("processing time of the latest message ... is multiplied by the number of in-process messages on that computer to produce a load index") and
- d) determining load distribution among the computers based on the load indexes ("values are calculated at all computers, and a computer with the smallest load index value is selected to process the message").

The APA does not explicitly disclose the additional limitations detailed below.

Sitbon teaches:

- b) determining estimated elongation rates of processing time for respective ones of the computers based on the estimated load states ("calculating the progress rate of the load", p3 14-15)
- c) calculating load Indexes of respective ones of the computers from the estimated elongation rates ("slope and average of the load calculated", p3 18-21) and

"means for choosing the least loaded server", p3 24-5.

It would have been obvious to combine Sitbon's teachings with the APA because the predictive slope computations increase accuracy of load indications, thereby improving the distributive processing efficacy. The APA as modified by Sitbon, however, does not explicitly disclose the additional limitations detailed below.

Hirata teaches "the time when a job is in this queue may be taken into consideration as a delay time due to the queue, to analyze a cause of delay of a job net", p16 20-39 which corresponds to the limitation of "including queuing time" in the processing measurements. It would have been obvious to combine Hirata's teachings with the APA as modified because the detailed analysis provides means to "effectively detect a causative location", p2 45-49 that is

integral in processing time computations, and thereby more accurately determine "loads of the computers" for the job distribution system.

As to claim 2, Sitbon (p5 32-39) teaches measuring load data at constant intervals and estimating an associated load state of the server.

As to claim 3, the APA teaches "the number of jobs at a computer" used to calculate processing metrics in a system, citing Kameda example.

As to claim 4, Sitbon (p5 32-39) teaches the ordered computations and resultant selection substantially as claimed.

As to claims 5-6, note the discussions of claims 2-3 supra.

As to claims 7-8, Sitbon teaches measuring load data at constant intervals and estimating associated load state for each of the servers, p5 32-39.

As to claim 11, Sitbon (p3 14-21) teaches the server progress evaluation as a metric for each of the systems.

As to claim 12, see the APA's teachings on load computations using in-process transaction indexes, and the periodic detection that calculates the load state with respect to the scheduled transactions.

As to claims 13-14, the APA teaches "processing time of the latest message ... is multiplied by the number of in-process messages on that computer to produce a load index value." These measurements are continuously updated to reflect system stats, and it would have naturally flowed from this prior art to use the updates (before/after) each job assignment to show system progress and load state.

As to claims 15-22 and 25-28, note the rejections of claims 1-8 and 11-14 above. Claims 15-22 and 25-28 are the same as claims 1-8 and 11-14, except claims 15-22 and 25-28 are apparatus claims and claims 1-8 and 11-14 are method claims.

As to claim 37, note the rejection of claim 1 above. Claim 37 is functionally equivalent to claim 1, but for the additional recitation that an elongation rate is a ratio of a processing time required for processing a transaction to a net processing time which is a sum of CPU time and an input/output time for processing the transaction. The APA and Sitbon, as discussed supra, clearly show that CPU time and I/O time have been used in combination such that it would have made obvious the estimated load states as claimed.

7. Claims 29-30, 32, 38-39 and 44-45 are rejected under 35 U.S.C. § 103(a) as being unpatentable over the APA, Sitbon, and Hirata as applied to claim 15, and further in view of Choquier et al. (U.S. Patent 5,774,668).

As to claim 29, Choquier (p3 29-48) teaches "Gateway microcomputers... determine the loads of the application servers ... and then passes the service request to the server having the greatest available CPU processing". It would have been obvious to combine Choquier's teachings with the APA as modified because the "service maps" provide an efficient means for updating the other nodes in the system with load info, thereby facilitating an improved load-balancing efficacy.

As to claims 30 and 32, Choquier (p10 39-45) teaches "every Gateway 126 receives and locally stores a copy of the new service map 136 to determine the states of the servers 120" for load monitoring and assigning. It would have been obvious to combine Choquier's teachings with the APA as modified because the "service maps" provide an efficient means for updating the other nodes in the system with load info, thereby facilitating an improved load-balancing efficacy.

As to claim 38, the APA teaches a load balancing method in a system ("Optimal Load Balancing in Distributed Computer Systems") comprising a plurality of computers (servers) for processing transaction processing requests ("distributing ... messages each requiring small-scale processing") originating from a plurality of terminals, comprising steps of:

estimating load states of respective ones of the computers ("processing time of the latest message ... is multiplied by the number of in-process messages on that computer to produce a load index") and

determining load distribution among the computers based on the load indexes ("values are calculated at all computers, and a computer with the smallest load index value is selected to process the message").

The APA does not explicitly disclose the additional limitations detailed below.

Sitbon teaches:

determining estimated elongation rates of processing time for respective ones of the computers based on the estimated load states ("calculating the progress rate of the load", p3 14-15)

calculating load Indexes of respective ones of the computers from the estimated elongation rates ("slope and average of the load calculated", p3 18-21) and "means for choosing the least loaded server", p3 24-5.

It would have been obvious to combine Sitbon's teachings with the APA because the predictive slope computations increase accuracy of load indications, thereby

improving the distributive processing efficacy. The APA as modified by Sitbon, however, does not explicitly disclose the additional limitations detailed below.

Hirata teaches "the time when a job is in this queue may be taken into consideration as a delay time due to the queue, to analyze a cause of delay of a job net", p16 20-39 which corresponds to the limitation of "including queuing time" in the processing measurements. It would have been obvious to combine Hirata's teachings with the APA as modified because the detailed analysis provides means to "effectively detect a causative location", p2 45-49 that is integral in processing time computations, and thereby more accurately determine "loads of the computers" for the job distribution system.

Choquier teaches "Gateway microcomputers... determine the loads of the application servers ... and then passes the service request to the server having the greatest available CPU processing", p3 29-48 which corresponds to the load balancing device and its position for managing the distribution operations. It would have been obvious to combine Choquier's teachings with the APA as modified because the "service maps" provide an efficient means for updating the other nodes in the system with load info, thereby facilitating an improved load-balancing efficacy.

As to claim 39, note the discussion of claim 38, with the addition of the "hot redirection technique" taught by Choquier, abstract for redistributing jobs so that they can be "transferred from one application server to another" to dynamically adjust request processing.

As to claims 44-45, note the rejections of claims 38-39 above. Claims 44-45 are the same as claims 38-39, except claims 44-45 are computer program product claims and claims 38-39 are method claims.

8. Claims 47-49 are rejected under 35 U.S.C. § 103(a) as being unpatentable over the APA in view of Sitbon et al. (U.S. Patent 5,951,634) and Tsuchiya et al. (U.S. Patent 5,978,844).

As to claim 47, the APA teaches a load balancing method in a system ("Optimal Load Balancing in Distributed Computer Systems") comprising a plurality of computers (servers) for processing transaction processing requests ("distributing ... messages each requiring small-scale processing") originating from a plurality of terminals, comprising steps of:

a) estimating load states of respective ones of the computers ("processing time of the latest message ... is multiplied by the number of in-process messages on that computer to produce a load index") and

d) determining load distribution among the computers based on the load indexes ("values are calculated at all computers, and a computer with the smallest load index value is selected to process the message").

The APA does not explicitly disclose the additional limitations detailed below.

Sitbon teaches:

- b) determining estimated elongation rates of processing time for respective ones of the computers based on the estimated load states ("calculating the progress rate of the load", p3 14-15)
- c) calculating load Indexes of respective ones of the computers from the estimated elongation rates ("slope and average of the load calculated", p3 18-21) and

"means for choosing the least loaded server", p3 24-5.

It would have been obvious to combine Sitbon's teachings with the APA because the predictive slope computations increase accuracy of load indications, thereby improving the distributive processing efficacy. The APA as modified by Sitbon, however, does not explicitly disclose the additional limitations detailed below.

Tsuchiya teaches "statistics units associated with respective processors for collecting the processing time . . . and (2) an adjusting unit for changing the allocation", p2 48-57 which corresponds to the correcting the estimated load states using current numbers of in-process transactions. It would have been obvious to combine Tsuchiya's teachings with the APA as modified because the statistical feedback (actual measurements) facilitate greater accuracy in adjusting assignments for the load balancing process.

As to claims 48-49, note the claim 47 discussion. Claims 48-49 are functionally equivalent to claim 47, and the respective limitations would likewise be obvious in view of the claim 47 discussion supra.

- 9. The prior art of record and not relied upon is considered pertinent to the applicant's disclosure. Specifically, the below reference(s) will also have relevancy to one or more elements of the Applicant's claimed invention as follows:
- U.S. Patent No. 6,587,938 to Eilert et al. which teaches the rate computations with elongation of the total I/O response time for jobs;
- U.S. Patent No. 6,377,975 to Florman which teaches the load indexing for determining optimal distributions;
- U.S. Patent No. 6,128,642 to Doraswamy et al. which teaches the processor utilization is calculated using actual job execution times;

U.S. Patent No. 6,006,248 to Nagae which teaches the cetral storage/control of load information and estimations; and,

U.S. Patent No. 5,881,284 to Kubo which teaches the job selection using evaluations of processing-load indexes.

10. Response to Applicant's Arguments:

During patent examination, the pending claims must be "given their broadest reasonable interpretation consistent with the specification." *In re Hyatt*, 211 F.3d 1367, 1372, 54 USPQ2d 1664, 1667 (Fed. Cir. 2000). Applicant always has the opportunity to amend the claims during prosecution, and broad interpretation by the examiner reduces the possibility that the claim, once issued, will be interpreted more broadly than is justified. *In re Prater*, 415 F.2d 1393, 1404-05, 162 USPQ 541, 550-51 (CCPA 1969)

With respect to claims 9-10, 23-24 and 31, the specified limitations have been carefully considered, and the presently presented 9-10, 23-24 and 31 would be allowable if rewritten in independent form *including all of the limitations of the base claim and the intervening claims*.

As to the rejected claims, the processing time measurements for balancing loads, in the manner recited, does not constitute a nonobvious improvement over the prior art.

In considering the load calculation recitations, it is noted that Applicant uses terminology that has broad meaning in the art, and thus requires a broad interpretation of the claims in determining patentability of the disclosed invention. Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. In re Van Geuns, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993). See also In re Zletz, 893 F.2d 319, 321-22, 13 USPQ2d 1320, 1322 (1989) "During patent examination the pending claims must be interpreted as broadly as their terms reasonably allow.... The reason is simply that during patent prosecution when claims can be amended, ambiguities should be recognized, scope and breadth of language explored, and clarification imposed.... An essential purpose of patent examination is to fashion claims that are precise, clear, correct, and unambiguous. Only in this way can uncertainties of claim scope be removed, as much as possible, during the administrative process."

Applicant's arguments, filed 30 August 2004, have been considered but are deemed moot in view of the new grounds of rejection necessitated by the amendments.

11. THIS ACTION IS MADE FINAL.

A SHORTENED STATUTORY PERIOD FOR RESPONSE TO THIS FINAL ACTION IS SET TO EXPIRE THREE MONTHS FROM THE DATE OF THIS ACTION. IN THE EVENT A FIRST RESPONSE IS FILED WITHIN TWO MONTHS OF THE MAILING DATE OF THIS FINAL ACTION AND THE ADVISORY ACTION IS NOT MAILED UNTIL AFTER THE END OF THE THREE-MONTH SHORTENED STATUTORY PERIOD, THEN THE SHORTENED STATUTORY PERIOD WILL EXPIRE ON THE DATE THE ADVISORY ACTION IS MAILED, AND ANY EXTENSION FEE PURSUANT TO 37 C.F.R. 1.136(a) WILL BE CALCULATED FROM THE MAILING DATE OF THE ADVISORY ACTION. IN NO EVENT WILL THE STATUTORY PERIOD FOR RESPONSE EXPIRE LATER THAN SIX MONTHS FROM THE DATE OF THIS FINAL ACTION.

12. Contact Information:

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system.

Status information for published applications may be obtained from either Private-PAIR or Public-PAIR.

Status information for unpublished applications is available through Private-PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov.

Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

All responses sent by U.S. Mail should be mailed to:

Commissioner for Patents

PO Box 1450

Alexandria, VA 22313-1450

Hand-delivered responses should be brought to Crystal Park Two, 2021 Crystal Drive, Arlington, VA., Sixth Floor (Receptionist). All hand-delivered responses will be handled and entered by the docketing personnel. Please do not hand deliver responses directly to the Examiner.

The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

All OFFICIAL faxes will be handled and entered by the docketing personnel. The date of entry will correspond to the actual FAX reception date unless that date is a Saturday, Sunday, or a Federal Holiday within the District of Columbia, in which case the official date of receipt will be the next business day. The application file will be promptly forwarded to the Examiner unless the application file must be sent to another area of the Office, e.g., Finance Division for fee charging, etc.

Any inquiry of a general nature or relating to the status of this application should be directed to the Group receptionist at (703) 305-9600.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to George Opie at (571) 272-3766 or via e-mail at *George.Opie@uspto.gov*. Internet e-mail should not be used where sensitive data will be exchanged or where there exists a possibility that sensitive data could be identified unless there is an express waiver of the confidentiality requirements under 35 U.S.C. 122 by the Applicant. Sensitive data includes confidential information related to patent applications.

MENG-AL T. AN SUPERVISORY PATENT EXAMINER

TECHNOLOGY CENTER 2100